International Journal of All Research Education and Scientific Methods (IJARESM),
ISSN: 2455-6211, Volume 13, Issue 5, May-2025, Available online at: www.ijaresm.com

Website Data Leakage Detection

Neeraj Appani[1], Shreyash Diwathe[2]

[1,2]Department of Computer Science, Faculty of Computer Science & Cyber Security
MIT-ADT University

----------**********----------

ABSTRACT

In the digital age, protecting sensitive information on websites has become increasingly important, and data breaches pose a serious threat to privacy, integrity, and business reputation. This paper describes a general technique for searching for information on the web using advanced analytics, anomaly detection algorithms, and machine learning. Our solutions focus on identifying unauthorized data, detecting vulnerabilities, and responding to breaches immediately. The planning process includes network analysis, behavioral analysis, and deep packet analysis to identify malicious patterns that indicate deletion information. We increase accuracy and reduce false positives by using AI-based predictive models. The goal is to enable organizations to maintain strong data protection, comply with privacy laws, and minimize cyber risks. Our results demonstrate the effectiveness of this approach in providing a large-scale, effective, and efficient solution for detecting and mitigating criminal networks.

Index Terms—Detection, Hooking, Signature-Based, Malware Rootkits, Anomaly Based, OS, API.

INTRODUCTION

One of the biggest takeaways from our 2023 event is that attackers are focused on evasion, which was also a major theme of M- Trends 2024. Their goal is to not have a search process (like last call and answer) and to protect the network for as long as possible by focusing on equipment, â€ œon the groundâ€ • and other technologies, using Zero-Day technology or no flaws in all-work security and other solutions. While attackers work hard to avoid detection, defenders continue to get better at detecting conflicts.

The average dwell time (the number of days an attacker stays in the system from the moment of attack until detection) continues in 2023 and is now 10 days (last year it was 16 days). This is a big win for the good guys, but itâ€ ™s still important to shorten the dwell time as ransomware will be detected faster. In addition, the Mandiant Red team typically completes its targets within 5 to 7 days, so defenders need to be vigilant. M-Trends 2024 provides readers with insights and other security measures, reports on financial attackersâ€ ™criminal and zero-day exploits, and specifically guides the behavior of the Chinese spy team. Other key takeaways from this report include: Changing phishing trends, such as attackersâ€ ™techniques for bypassing multiple authentication methods using social media, SMS, and other communication technologies such as adversaries and other strategies Cloud intrusion trends, such as attackers targeting cloud infrastructure and cloud technology. Share with the Red and Purple team how smart, focused new technologies can help organizations achieve better outcomes. Mandiant consultants are on the front lines, investigating, analyzing, and learning about the latest cyberattacks. Consultants monitor clients for new challenges, systems, and processes, and help them remediate, change, and learn. By publishing our M-Trends report each year, we strive to provide important information to those working in defense organizations and thus share what we have learned with the security community. The information in this report has been sanitized to protect the privacy of victims and their information.

Internal detection is when an organization detects that it has been compromised, either through an alert on a security device or an internal activity alert.

An external alert is when an external organization, such as a law enforcement agency, cybersecurity company, or business partner, notifies an organization that a threat has been detected. In some cases, attackers will provide this alert through means such as a ransom note.

Detection By Source

Internal detection is when an organization detects that it has been compromised, either through notifications from security

devices or internal activity reports. or when a business partner notifies the organization that its security has been compromised. In some cases, attackers make this notification through means such as a ransom note. source of the incident, 46% first saw evidence of internal involvement. However, isolated ransomware-related intrusions suggest that it is more common for organizations to learn about ransomware-related incidents from other sources. In order to combat ransomware, 70% of organizations receive external alerts, often demanding ransoms from attackers. For non- ransomware intrusions, the ratio of internal detection to external detection is 50-50. 85% of intrusions detected internally do not involve ransomware. many. However, Mandiant saw a nine percent decrease in external reports. This change from last year, along with the high level of agreement found in non-ransomware cases, suggests that organizations are getting better at detecting malicious behavior on their networks.

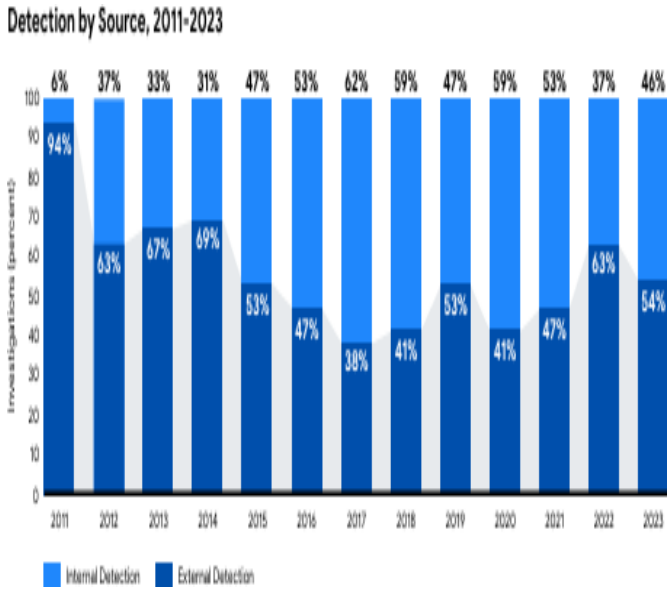


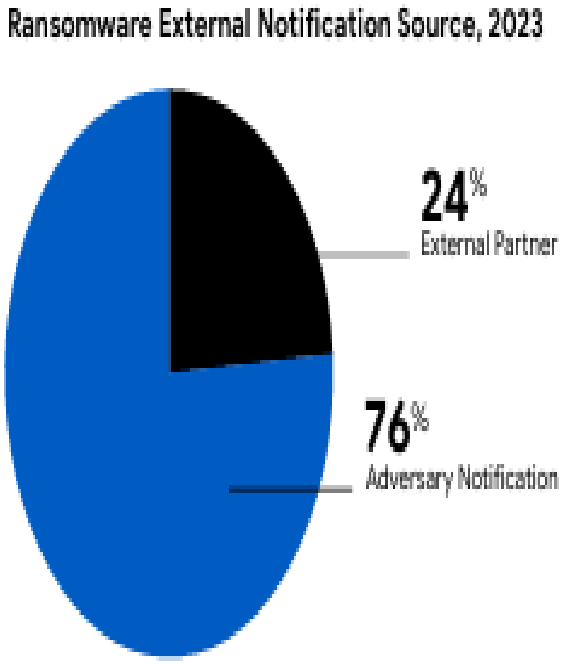
Fig.3: Detection by source 2011- 2023

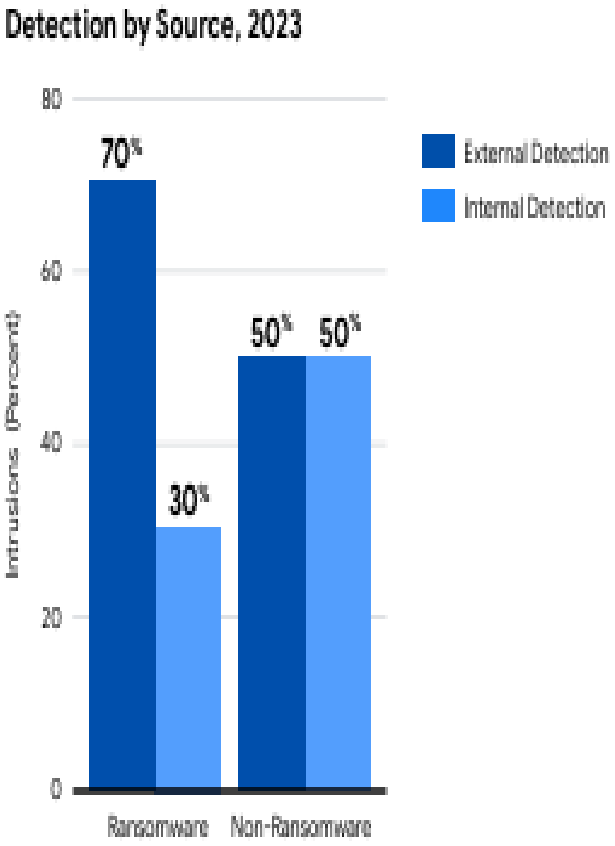
Ransomware-related intrusion provides attackers with access or interaction, the main purpose of which is to understand the information and extract money from the target, prevent further destruction of the damage or eliminate the evil.

Ransomeware Related Intrutions

In 70% of cases, organizations investigated ransomware-related attacks from other sources. In three-quarters of breaches, organizations experienced a ransomware incident reported by an attacker. This is based on the ransomware business model, where attackers intentionally notify organizations about ransomware crimes out of the blue and demand payment.

Three-quarters of external reports of ransomware crimes come from external partners, such as law enforcement or security companies. In 2022, two-thirds of external reports of ransomware crimes were attack reports, while one-third came from external actors.


Ransomware External Notification Source, 2023
24%
External Partner
76%
Adversary Notification


Detection by Source, 2023
80
60
40
20
0
Intrusions (Percent)
70%
30%
50%
50%
External Detection
Internal Detection
Ransomware
Non-Ransomware

The dwell time is calculated as the number of days an attacker spends in the affected area before being detected. The average represents the value of a point in the dataset relative to the dimension

Dwell Time

The average time between the world continues to decrease, marking another minimum intrusion time for the first time and discovered during the entire introduction of M-Trends. By 2023, most organizations will detect intrusions within 10 days of the

METHODOLOGY

In recent years, the main purpose of information security has been threatened by the organization's employees, partners and partners whose information is confidential. One of the most important measures proposed in the literature is data access control, which allows limiting operations that go beyond access control and allow the leakage of confidential information and control its use. Threats from employees to confidential information reinforce the importance of the security of confidential information management; To solve initial breach. This is down to almost a week compared to 16 days this problem, they proposed a confidential information security in 2022. Mandiant Defenders saw a significant improvement in the model with distributed control. The first of these is based on the global average duration of all alerts in 2023. The global average duration of the external alert period decreased from 19 days in 2022 to 13 days in 2023 due to the shortest duration. Another possible explanation for this decrease could be an increase in ransomware attack alerts. Continuing the trend, when defenders see an enemy attacking from the inside, they act faster than the time in the middle of the whole. The global median time between access detections was 9 days in 2023, compared to 13 days in 2022 and 18 days in 2021.

Global Median Dwell Time, 2011-2023

	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023
All	416	243	229	205	146	99	101	78	56	24	21	16	10
External	–	–	–	–	320	107	186	184	141	73	28	19	13
Internal	–	–	–	–	56	80	57.5	50.5	30	12	18	13	9

Attack Mission

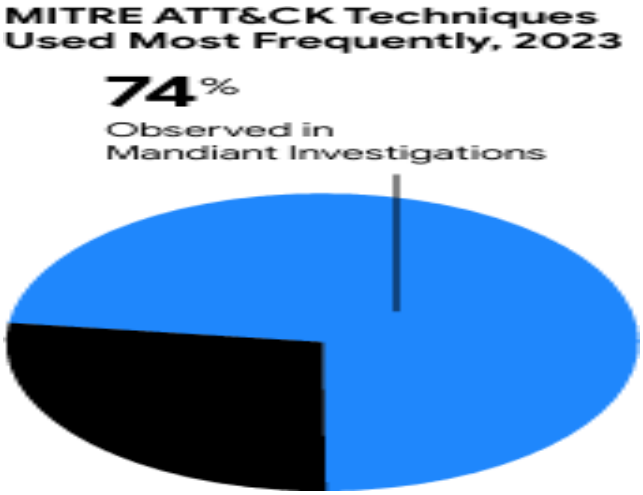
First Virus Vectors In 2023, Mandiant experts once again found a flaw that was used as the biggest virus for the adversary. 38% of the attacks where the vector was detected started negatively. This is a 6 point increase from 2022 and follows what defenders experienced in 2021. Phishing is still the second most common intrusion vector. But that number will drop in 2023, from 17% of penetrations to 22% in 2022. Security. Past impacts, the third most effective vector used by attackers in 2023. Visit begins.

This growth may be related to the ongoing collaboration of various malware vendors with the ransomware ecosystem and ransomware affiliates.

Storage of the content in encrypted form using the symmetric encryption algorithm to ensure that only authorized users can decrypt the content and store information management, and allows the user to understand the level of the right to use confidential information and write all the information. User work: Also use the hash function to provide the content. To control the sharing of confidential information, use the client-server policy that the user cannot share confidential information without permission from the server, use administrator rights and install authorized servers on the user's computer to view. Control the actions of users and prohibit unauthorized actions.

A DLPS based on control usage and dynamic data monitoring (OCAWin) is proposed in the research [55]. The system can monitor the operation calls made to Windows APIs to prevent or modify data flows that pose a threat to confidential data. The solution exposes the open functions of the middleware to ensure that the data remains in memory and also establishes access control rules such as read-only, same, send-out, registration, backup and authorization. For access control, a shared authentication and key communication method is established between the user and the server, using the SM2 algorithm for control. By capturing the biometric identifiers of the users related to the product (data), the data associated with the content

entered by the user's biometric data is not stored, but the quality of the interaction of the two is preserved. In this way, in the event of a loss, organizations will be able to identify the user who accessed the data and the risk of biometric data being compromised will be reduced. The mini-filter controls the use of confidential information by controlling OS 1/0 functions. The system plans to block IO operations from an external storage device. Additionally, the idea of â€ ‹ â€ ‹ restricting the blacklisting of classified information and preventing further attempts at the process was also adopted.



CONCLUSION AND FUTURE RESEARCH

With the increasing number of threats in todayâ€ ™s digital environment, detecting data breaches is an important aspect of website security. Websites often contain sensitive information such as personal information, financial details, and proprietary and confidential business information. In todayâ€ ™s connected digital world, protecting sensitive information is more important than ever. As businesses and websites continue to handle confidential information, the threat of data breaches and leaks is also increasing.

REFERENCES

[1]. Data breach quick view. 2015. Data breach trends. Available at https://www.riskbasedsecurity.com/2015-data-breach- quickview/. Accessed on 5 Sept 2019.
[2]. Data leakage news. Available
[3]. at https://infowatch.com/analytics/leaks_monitoring. Accessed on 5 Sept 2019.
[4]. IBM security's cost of a data breach report 2019. Available
[5]. at https://www.ibm.com/security/data-breach. Accessed on 5 Sept 2019.
[6]. Asaf, Shabtai, Yuval Elovici, and Lior Rokach. 2012. A survey of data leakage detection and prevention solutions, 1st ed. Springer: New York Heidelberg Dordrecht
[7]. London. https://doi.org/10.1007/978-1-4614-2053-8.